ICDM 2010

The 10th IEEE International Conference on Data Mining

14-17 December 2010 Sydney, Australia

Edited by Geoffrey I. Webb, Bing Liu, Chengqi Zhang, Dimitrios Gunopulos, and Xindong Wu





Conference Information

Papers by Session

Papers by Author

Search

Getting Started

Trademarks

Published by:





			T	
Report Docume	entation Page			Form Approved MB No. 0704-0188
Public reporting burden for the collection of information is estimated maintaining the data needed, and completing and reviewing the collection cluding suggestions for reducing this burden, to Washington Header VA 22202-4302. Respondents should be aware that notwithstanding does not display a currently valid OMB control number.	ction of information. Send comments uarters Services, Directorate for Info	regarding this burden estimat rmation Operations and Repo	te or any other aspect of rts, 1215 Jefferson Davi	this collection of information, is Highway, Suite 1204, Arlington
1. REPORT DATE 22 DEC 2010	2. REPORT TYPE		3. DATES COVE 14-12-2010	RED to 17-12-2010
4. TITLE AND SUBTITLE The IEEE International Conference on Data Mining (10th)		n) Held in	5a. CONTRACT FA2386-10-	
Sydney, Australia on 14-17 December	2010		5b. GRANT NUM	/BER
			5c. PROGRAM E	LEMENT NUMBER
6. AUTHOR(S)			5d. PROJECT NU	JMBER
			5e. TASK NUMB	ER
			5f. WORK UNIT	NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Technology, Sydney,PO Box 123,Broadway, 2007,Sydney,AU,00000		NSW	8. PERFORMING NUMBER ; CSP-10104	G ORGANIZATION REPORT
9. SPONSORING/MONITORING AGENCY NAME(S) AOARD, Unit 45002, APO, AP, 96338	` '		10. SPONSOR/M AOARD	ONITOR'S ACRONYM(S)
			11. SPONSOR/M NUMBER(S) CSP-101048	ONITOR'S REPORT
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribute	tion unlimited			
13. SUPPLEMENTARY NOTES See ADM002348. Supporting docume government purpose rights.	nts are attached to t	he report as sepa	arate files (PI	OF). U.S.
14. ABSTRACT The proceedings of the conference is i including algorithms, software and sy machine learning, pattern recognition knowledge-based systems to high perf (3), regular papers (72), short papers	stems, and applicati , databases and data ormance computing	ons. Thus the top a warehousing, d g. This proceedin	pics ranges fr lata visualiza	om statistics, tion,
15. SUBJECT TERMS Data Mining, Machine Learning, Arti	ficial Intelligence, K	nowledge Acqui	sition, Know	ledge Discovery
16. SECURITY CLASSIFICATION OF:		17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON

Public

Release

138

c. THIS PAGE

unclassified

a. REPORT

unclassified

b. ABSTRACT

unclassified

Conference Information

2010 IEEE International Conference on Data Mining

- Welcome Message from the Conference Chairs
- Message from the Program Committee Co-Chairs
- Organizing Committee
- □ Program Committee
- ☐ Title Page (Book version)
- ☐ Copyright Page (Book version)
- Table of Contents (Book version)
- Author Index (Book version)
- Publisher Information (Book version)







Sessions

- □ Keynote Abstracts
- □ Regular Papers
- Short Papers
- Tutorials







Keynote Abstracts

- Mining Billion-node Graphs: Patterns, Generators and Tools Christos Faloutsos
- Assessing the Significance of Groups in High-Dimensional Data Geoff McLachlan
- ☐ 10 Years of Data Mining Research: Retrospect and Prospect Xindong Wu







Regular Papers

- Detecting Novel Discrepancies in Communication Networks James Abello, Tina Eliassi-Rad, and Nishchal Devanur
- Multi-agent Random Walks for Local Clustering on Graphs Morteza Alamgir and Ulrike von Luxburg
- Spatiotemporal Event Detection in Mobility Network Tom S. Au, Rong Duan, Heeyoung Kim, and Guang-Qin Ma
- □ An Unsupervised Approach to Modeling Personalized Contexts of Mobile Users Tengfei Bao, Happia Cao, Enhong Chen, Jilei Tian, and Hui Xiong
- ☐ Fast and Flexible Multivariate Time Series Subsequence Search Kanishka Bhaduri, Qiang Zhu, Nikunj C. Oza, and Ashok N. Srivastava
- ☐ iSAX 2.0: Indexing and Mining One Billion Time Series
 Alessandro Camerra, Themis Palpanas, Jin Shieh, and Eamonn Keogh
- □ Abstraction Augmented Markov Models
 Cornelia Caragea, Adrian Silvescu, Doina Caragea, and Vasant Honavar
- □ A Graph-Based Approach for Multi-folder Email Classification Sharma Chakravarthy, Aravind Venkatachalam, and Aditya Telang







Scalable Influence Maximization in Social Networks under the Linear Threshold Model Wei Chen, Yifei Yuan, and Li Zhang CLUSMASTER: A Clustering Approach for Sampling Data Streams in Sensor **Networks** Alzennyr Da Silva, Raja Chiky, and Georges Hebrail Bayesian Maximum Margin Clustering Bo Dai, Baogang Hu, and Gang Niu Viral Marketing for Multiple Products Samik Datta, Anirban Majumder, and Nisheeth Shrivastava Finding Local Anomalies in Very High Dimensional Space Timothy de Vries, Sanjay Chawla, and Michael E. Houle PGLCM: Efficient Parallel Mining of Closed Frequent Gradual Itemsets Trong Dinh Thac Do, Anne Laurent, and Alexandre Termier Sequential Latent Dirichlet Allocation: Discover Underlying Topic Structures within a Document

Lan Du, Wray Lindsay Buntine, and Huidong Jin







Subgroup Discovery Meets Bayesian Networks—An Exceptional Model Mining Approach Wouter Duivesteijn, Arno Knobbe, Ad Feelders, and Matthijs van Leeuwen
Feature Selection for Unsupervised Learning Using Random Cluster Ensembles Haytham Elghazel and Alex Aussem
Learning Attribute-to-Feature Mappings for Cold-Start Recommendations Zeno Gantner, Lucas Drumond, Christoph Freudenthaler, Steffen Rendle, and Lars Schmidt-Thieme
An Extensive Empirical Study on Semi-supervised Learning Yuanyuan Guo, Xiaoda Niu, and Harry Zhang
Efficient Discovery of the Top-K Optimal Dependency Rules with Fisher's Exact Test of Significance Wilhelmiina Hämäläinen
A Variance Reduction Framework for Stable Feature Selection Yue Han and Lei Yu
Exponential Family Tensor Factorization for Missing-Values Prediction and Anomaly Detection Kohei Hayashi, Takashi Takenouchi, Tomohiro Shibata, Yuki Kamiya, Daishi Kato, Kazuo Kunieda, Keiji Yamada, and Kazushi Ikeda







Rare Category Characterization Jingrui He, Hanghang Tong, and Jaime Carbonell
Algorithm for Discovering Low-Variance 3-Clusters from Real-Valued Datasets Zhen Hu and Raj Bhatnagar
Improved Consistent Sampling, Weighted Minhash and L1 Sketching Sergey loffe
An Approach Based on Tree Kernels for Opinion Mining of Online Product Reviews Peng Jiang, Chunxia Zhang, Hongping Fu, Zhendong Niu, and Qing Yang
A Pairwise-Systematic Microaggregation for Statistical Disclosure Control Md. Enamul Kabir, Hua Wang, and Yanchun Zhang
Multi-label Feature Selection for Graph Classification Xiangnan Kong and Philip S. Yu
A Binary Decision Diagram-Based One-Class Classifier Takuro Kutsuna
Detecting Blackhole and Volcano Patterns in Directed Networks Zhongmou Li, Hui Xiong, Yanchi Liu, and Aoying Zhou
Exploiting Local Data Uncertainty to Boost Global Outlier Detection Bo Liu, Jie Yin, Yanshan Xiao, Longbing Cao, and Philip S. Yu







Training Conditional Random Fields Using Transfer Learning for Gesture Recognition Jie Liu, Kai Yu, Yi Zhang, and Yalou Huang
Stratified Sampling for Data Mining on the Deep Web Tantan Liu, Fan Wang, and Gagan Agrawal
Learning Markov Network Structure with Decision Trees Daniel Lowd and Jesse Davis
Towards Structural Sparsity: An Explicit I2/I0 Approach Dijun Luo, Chris Ding, and Heng Huang
Multi-document Summarization Using Minimum Distortion Tengfei Ma and Xiaojun Wan
A Log-Linear Model with Latent Features for Dyadic Prediction Aditya Krishna Menon and Charles Elkan
Edge Weight Regularization over Multiple Graphs for Similarity Learning Pradeep Muthukrishnan, Dragomir Radev, and Qiaozhu Mei
A New SVM Approach to Multi-instance Multi-label Learning Nam Nguyen
Bayesian Aggregation of Binary Classifiers







Permutations as Angular Data: Efficient Inference in Factorial Spaces Sergey M. Plis, Terran Lane, and Vince D. Calhoun Separation of Interleaved Web Sessions with Heuristic Search Marko Poženel, Viljan Mahnic, and Matjaž Kukar Consequences of Variability in Classifier Performance Estimates Troy Raeder, T. Ryan Hoens, and Nitesh V. Chawla Mining Sensor Streams for Discovering Human Activity Patterns over Time Parisa Rashidi and Diane J. Cook Decision Trees for Uplift Modeling Piotr Rzepakowski and Szymon Jaroszewicz Co-clustering of Lagged Data Eran Shaham, David Sarne, and Boaz Ben-Moshe Polishing the Right Apple: Anytime Classification Also Benefits Data Streams with Constant Arrival Times Jin Shieh and Eamonn Keogh Discovering Correlated Subspace Clusters in 3D Continuous-Valued Data

Kelvin Sim, Zeyar Aung, and Vivekanand Gopalkrishnan







gSkeletonClu: Density-Based Network Clustering via Structure-Connected Tree Division or Agglomeration Heli Sun, Jianbin Huang, Jiawei Han, Hongbo Deng, Peixiang Zhao, and Boqin Feng
LogTree: A Framework for Generating System Events from Raw Textual Logs Liang Tang and Tao Li
Mining Closed Strict Episodes Nikolaj Tatti and Boris Cule
Multi-dimensional Mass Estimation and Mass-based Clustering Kai Ming Ting and Jonathan R. Wells
minCEntropy: A Novel Information Theoretic Approach for the Generation of Alternative Clusterings Nguyen Xuan Vinh and Julien Epps
A Conscience On-line Learning Approach for Kernel-Based Clustering Chang-Dong Wang, Jian-Huang Lai, and Jun-Yong Zhu
Weighted Feature Subset Non-negative Matrix Factorization and Its Applications to Document Understanding Dingding Wang, Tao Li, and Chris Ding
Learning a Bi-Stochastic Data Similarity Matrix Fei Wang, Ping Li, and Arnd Christian König







Active Spectral Clustering Xiang Wang and Ian Davidson
Discovering Overlapping Groups in Social Media Xufei Wang, Lei Tang, Huiji Gao, and Huan Liu
Adaptive Distances on Sets of Vectors Adam Woznica and Alexandros Kalousis
SMILE: A Similarity-Based Approach for Multiple Instance Learning Yanshan Xiao, Bo Liu, Longbing Cao, Jie Yin, and Xindong Wu
Modeling Information Diffusion in Implicit Networks Jaewon Yang and Jure Leskovec
Term Filtering with Bounded Error Zi Yang, Wei Li, Jie Tang, and Juanzi Li
Exploiting Unlabeled Data to Enhance Ensemble Diversity Min-Ling Zhang and Zhi-Hua Zhou
Constraint Based Dimension Correlation and Distance Divergence for Clustering High-Dimensional Data Xianchao Zhang, Yao Wu, and Yang Qiu
Active Learning from Multiple Noisy Labelers with Varied Costs Yaling Zheng, Stephen Scott, and Kun Deng







A Novel Contrast Co-learning Framework for Generating High Quality Training Data Zeyu Zheng, Jun Yan, Shuicheng Yan, Ning Liu, Zheng Chen, and Ming Zhang Network Simplification with Minimal Loss of Connectivity Fang Zhou, Sébastien Malher, and Hannu Toivonen Improving Kernel Methods through Complex Data Mapping Hang Zhou, Fabio Ramos, and Eric Nettleton NESVM: A Fast Gradient Method for Support Vector Machines Tianyi Zhou, Dacheng Tao, and Xindong Wu Clustering Large Attributed Graphs: An Efficient Incremental Approach Yang Zhou, Hong Cheng, and Jeffrey Xu Yu Mother Fugger: Mining Historical Manuscripts with Local Color Patches Qiang Zhu and Eamonn Keogh D-LDA: A Topic Modeling Approach without Constraint Generation for Semi-defined Classification Fuzhen Zhuang, Ping Luo, Zhiyong Shen, Qing He, Yuhong Xiong, and Zhongzhi Shi







Short Papers

- SONNET: Efficient Approximate Nearest Neighbor Using Multi-core Mohammad Al Hasan, Hilmi Yildirim, and Abhirup Chakraborty
- Two of a Kind or the Ratings Game? Adaptive Pairwise Preferences and Latent Factor Models

Suhrid Balakrishnan and Sumit Chopra

- Document Similarity Self-Join with MapReduce
 Ranieri Baraglia, Gianmarco De Francisci Morales, and Claudio Lucchese
- Quantification via Probability Estimators
 Antonio Bella, Cèsar Ferri, José Hernández-Orallo, and María José Ramírez-Quintana
- Learning Collaborative Filtering and Its Application to People to People Recommendation in Social Networks

Xiongcai Cai, Michael Bain, Alfred Krzywicki, Wayne Wobcke, Yang Sok Kim, Paul Compton, and Ashesh Mahidadia

- Approximation of Frequentness Probability of Itemsets in Uncertain Data Toon Calders, Calin Garboni, and Bart Goethals
- On Finding Frequent Patterns in Event Sequences
 Andrea Campagna and Rasmus Pagh







Active Improvement of Hierarchical Object Features under Budget Constraints Nicolas Cebron Pseudo Conditional Random Fields: Joint Training Approach to Segmenting and Labeling Sequence Data Shing-Kit Chan and Wai Lam Location and Scatter Matching for Dataset Shift in Text Mining Bo Chen, Wai Lam, Ivor Tsang, and Tak-Lam Wong Learning Preferences with Millions of Parameters by Enforcing Sparsity Xi Chen, Bing Bai, Yanjun Qi, Qihang Lin, and Jaime Carbonell QMAS: Querying, Mining and Summarization of Multi-modal Databases Robson L.F. Cordeiro, Fan Guo, Donna S. Haverkamp, James H. Horne, Ellen K. Hughes, Gunhee Kim, Agma J.M. Traina, Caetano Traina Jr., and Christos Faloutsos Block-GP: Scalable Gaussian Process Regression for Multimodal Data Kamalika Das and Ashok N. Srivastava Active Learning with Human-Like Noisy Oracle Jun Du and Charles X. Ling Monotone Relabeling in Ordinal Classification



Ad Feelders





The Effect of History on Modeling Systems' Performance: The Problem of the Demanding Lord George Giannakopoulos and Themis Palpanas
Resilient K-d Trees: K-Means in Space Revisited Fabian Gieseke, Gabriel Moruz, and Jan Vahrenhold
Advertising Campaigns Management: Should We Be Greedy? Sertan Girgin, Jeremie Mary, Philippe Preux, and Olivier Nicol
Accelerating Radius-Margin Parameter Selection for SVMs Using Geometric Bounds Ben Goodrich, David Albrecht, and Peter Tischer
Enhancing Single-Objective Projective Clustering Ensembles Francesco Gullo, Carlotta Domeniconi, and Andrea Tagarelli
Minimizing the Variance of Cluster Mixture Models for Clustering Uncertain Objects Francesco Gullo, Giovanni Ponti, and Andrea Tagarelli
Subspace Clustering Meets Dense Subgraph Mining: A Synthesis of Two Paradigms Stephan Günnemann, Ines Färber, Brigitte Boden, and Thomas Seidl







Multi-stream Join Answering for Mining Significant Cross-Stream Correlations Robert Gwadera
Category Mining by Heterogeneous Data Fusion Using PdLSI Model in a Retail Service Tsukasa Ishigaki, Takeshi Takenaka, and Yoichi Motomura
Content-Based Methods for Predicting Web-Site Demographic Attributes Santosh Kabbur, Eui-Hong Han, and George Karypis
Discrimination Aware Decision Tree Learning Faisal Kamiran, Toon Calders, and Mykola Pechenizkiy
Patterns on the Connected Components of Terabyte-Scale Graphs U. Kang, Mary McGlohon, Leman Akoglu, and Christos Faloutsos
Attribution of Conversion Events to Multi-channel Media Brendan Kitts, Liang Wei, Dyng Au, Amanda Powter, and Brian Burdick
Mining Public Transport Usage for Personalised Intelligent Transport Systems Neal Lathia, Jon Froehlich, and Licia Capra
Micro-blogging Sentiment Detection by Collaborative Online Learning Guangxia Li, Steven C.H. Hoi, Kuiyu Chang, and Ramesh Jain
Enforcing Vocabulary <i>k</i> -Anonymity by Semantic Similarity Based Clustering Jungiang Liu and Ke Wang







Efficient Probabilistic Latent Semantic Analysis with Sparsity Control Sen Liu, Chaolun Xia, and Xiaohong Jiang
Understanding of Internal Clustering Validation Measures Yanchi Liu, Zhongmou Li, Hui Xiong, Xuedong Gao, and Junjie Wu
Transfer Learning via Cluster Correspondence Inference Mingsheng Long, Wei Cheng, Xiaoming Jin, Jianmin Wang, and Dou Shen
Supervised Link Prediction Using Multiple Sources Zhengdong Lu, Berkant Savas, Wei Tang, and Inderjit S. Dhillon
Addressing Concept-Evolution in Concept-Drifting Data Streams Mohammad M. Masud, Qing Chen, Latifur Khan, Charu Aggarwal, Jing Gao, Jiawei Han, and Bhavani Thuraisingham
Sparse Boolean Matrix Factorizations Pauli Miettinen
On the Computation of Stochastic Search Variable Selection in Linear Regression with UDFs Mario Navas, Carlos Ordonez, and Veerabhadran Baladandayuthapani
Data Editing Techniques to Allow the Application of Distance-Based Outlier Detection to Streams Vit Niennattrakul, Eamonn Keogh, and Chotirat Ann Ratanamahatana







Anomaly Detection Using an Ensemble of Feature Models Keith Noto, Carly Brodley, and Donna Slonim Assessing Data Mining Results on Matrices with Randomization Markus Ojala A Generalized Linear Threshold Model for Multiple Cascades Nishith Pathak, Arindam Banerjee, and Jaideep Srivastava Recommending Social Events from Mobile Phone Location Data Daniele Quercia, Neal Lathia, Francesco Calabrese, Giusy Di Lorenzo, and Jon Crowcroft On Normalizing Fuzzy Coincidence Matrices to Compare Fuzzy and/or Possibilistic Partitions with the Rand Index R. Quèré, H. Le Capitaine, N. Fraisseix, and C. Frélicot Financial Forecasting with Gompertz Multiple Kernel Learning Han Qin, Dejing Dou, and Yue Fang Leveraging D-Separation for Relational Data Sets Matthew J.H. Rattigan and David Jensen **Factorization Machines**



Steffen Rendle





Accelerating Dynamic Time Warping Subsequence Search with GPUs and FPGAs Doruk Sart, Abdullah Mueen, Walid Najjar, Eamonn Keogh, and Vit Niennattrakul
An Approach for Automatic Sleep Stage Scoring and Apnea-Hypopnea Detection Tim Schlüter and Stefan Conrad
Bonsai: Growing Interesting Small Trees Stephan Seufert, Srikanta Bedathur, Julian Mestre, and Gerhard Weikum
Mixed-Membership Stochastic Block-Models for Transactional Networks Mahdi Shafiei and Hugh Chipman
Generalized Probabilistic Matrix Factorizations for Collaborative Filtering Hanhuai Shan and Arindam Banerjee
Topic Modeling Ensembles Zhiyong Shen, Ping Luo, Shengwen Yang, and Xukun Shen
Interval-valued Matrix Factorization with Applications Zhiyong Shen, Liang Du, Xukun Shen, and Yidong Shen
Efficient Semi-supervised Spectral Co-clustering with Constraints Xiaoxiao Shi, Wei Fan, and Philip S. Yu
Transfer Learning on Heterogenous Feature Spaces via Spectral Transformation Xiaoxiao Shi, Qi Liu, Wei Fan, Philip S. Yu, and Ruixin Zhu







One-Class Matrix Completion with Low-Density Factorizations Vikas Sindhwani, Serhat S. Bucak, Jianying Hu, and Aleksandra Mojsilovic A System for Mining Temporal Physiological Data Streams for Advanced **Prognostic Decision Support** Jimeng Sun, Daby Sow, Jianying Hu, and Shahram Ebadollahi Averaged Stochastic Gradient Descent with Feedback: An Accurate, Robust, and Fast Training Method Xu Sun, Hisashi Kashima, Takuya Matsuzaki, and Naonori Ueda Visualizing Graphs Using Minimum Spanning Dendrograms Daniel Svonava and Michail Vlachos Tru-Alarm: Trustworthiness Analysis of Sensor Networks in Cyber-Physical Systems Lu-An Tang, Xiao Yu, Sangkyum Kim, Jiawei Han, Chih-Chieh Hung, and Wen-Chih Peng Node Similarities from Spreading Activation Kilian Thiel and Michael R. Berthold On the Vulnerability of Large Graphs Hanghang Tong, B. Aditya Prakash, Charalampos Tsourakakis, Tina Eliassi-Rad, Christos Faloutsos, and Duen Horng Chau







Testing the Significance of Patterns in Data with Cluster Structure Niko Vuokko and Petteri Kaski
Compressed Nonnegative Sparse Coding Fei Wang and Ping Li
Anonymizing Temporal Data Ke Wang, Yabo Xu, Raymond Chi-Wing Wong, and Ada Wai-Chee Fu
Homotopy Regularization for Boosting Zheng Wang, Yangqiu Song, and Changshui Zhang
What Do People Want in Microblogs? Measuring Interestingness of Hashtags in <i>Twitte</i> r Jianshu Weng, Ee-Peng Lim, Qi He, and Cane Wing-Ki Leung
Probabilistic Inference Protection on Anonymized Data Raymond Chi-Wing Wong, Ada Wai-Chee Fu, Ke Wang, Yabo Xu, Jian Pei, and Philip S. Yu
Collaborative Learning between Visual Content and Hidden Semantic for Image Retrieval Jun Wu, Ming-Yu Lu, and Chun-Li Wang
Max-Clique: A Top-Down Graph-Based Approach to Frequent Pattern Mining Yan Xie and Philip S. Yu







Personalizing Web Page Recommendation via Collaborative Filtering and Topic-Aware Markov Model Qingyan Yang, Ju Fan, Jianyong Wang, and Lizhu Zhou
Passive Sampling for Regression Hwanjo Yu and Sungchul Kim
Modeling Experts and Novices in Citizen Science Data for Species Distribution Modeling Jun Yu, Weng-Keen Wong, and Rebecca A. Hutchinson
Causal Discovery from Streaming Features Kui Yu, Xindong Wu, Hao Wang, and Wei Ding
ABS: The Anti Bouncing Model for Usage Data Streams Chongsheng Zhang, Florent Masseglia, and Yves Lechevallier
Classifier and Cluster Ensembles for Mining Concept Drifting Data Streams Peng Zhang, Xingquan Zhu, Jianlong Tan, and Li Guo
Graph-Based Semi-supervised Learning with Adaptive Similarity Estimation Xianchao Zhang, Yansheng Jiang, Wenxin Liang, and Xin Han
K-AP: Generating Specified K Clusters by Efficient Affinity Propagation Xiangliang Zhang, Wei Wang, Kietil Nørvåg, and Michèle Sebag







- MoodCast: Emotion Prediction via Dynamic Continuous Factor Graph Model Yuan Zhang, Jie Tang, Jimeng Sun, Yiran Chen, and Jinghai Rao
- Hierarchical Ensemble Clustering
 Li Zheng, Tao Li, and Chris Ding
- ☐ Frequent Instruction Sequential Pattern Mining in Hardware Sample Data Jia Zou, Jing Xiao, Rui Hou, and Yanqi Wang
- ☐ Efficient Episode Mining with Minimal and Non-overlapping Occurrences Huisheng Zhu, Peng Wang, Xianmang He, Yujia Li, Wei Wang, and Baile Shi







Tutorials

- □ Spatial and Spatio-temporal Data Mining Vania Bogorny and Shashi Shekhar
- □ Knowledge Discovery in Academic Drug Discovery Programs: Opportunities and Challenges
 □ Jun Huan
- How to Do Good Data Mining Research and Get it Published in Top Venues Eamonn Keogh
- Discovering Multiple Clustering Solutions: Grouping Objects in Different Views of the Data

Emmanuel Müller, Stephan Günnemann, Ines Färber, and Thomas Seidl







A	☐ Baladandayuthapani, Veerabhadran
	□ Balakrishnan, Suhrid
☐ Abello, James	■ Banerjee, Arindam
□ Aggarwal, Charu	■ Bao, Tengfei
☐ Agrawal, Gagan	☐ Baraglia, Ranieri
□ Akoglu, Leman	☐ Bedathur, Srikanta
□ Al Hasan, Mohammad	☐ Bella, Antonio
□ Alamgir, Morteza	☐ Ben-Moshe, Boaz
□ Albrecht, David	☐ Berthold, Michael R.
☐ Au, Dyng	☐ Bhaduri, Kanishka
☐ Au, Tom S.	☐ Bhatnagar, Raj
□ Aung, Zeyar	☐ Boden, Brigitte
☐ Aussem, Alex	☐ Bogorny, Vania
В	□ Brodley, Carly
<u> </u>	■ Bucak, Serhat S.
☐ Bai, Bing	■ Buntine, Wray Lindsay
■ Bain, Michael	







☐ Burdick, Brian	Chakravarthy, Sharma
	Chan, Shing-Kit
<u>C</u>	Chang, Kuiyu
☐ Cai, Xiongcai	Chau, Duen Horng
☐ Calabrese, Francesco	Chawla, Nitesh V.
☐ Calders, Toon	Chawla, Sanjay
☐ Calhoun, Vince D.	Chen, Bo
☐ Camerra, Alessandro	Chen, Enhong
☐ Campagna, Andrea	Chen, Qing
☐ Cao, Happia	Chen, Wei
☐ Cao, Longbing	Chen, Xi
☐ Capra, Licia	Chen, Yiran
☐ Caragea, Cornelia	Chen, Zheng
☐ Caragea, Doina	Cheng, Hong
☐ Carbonell, Jaime	Cheng, Wei
☐ Cebron, Nicolas	Chiky, Raja
☐ Chakraborty Abbirup	







☐ Chipman, Hugh	□ Datta, Samik
☐ Choi, Seungjin	Davidson, lan
☐ Chopra, Sumit	□ Davis, Jesse
□ Compton, Paul	□ De Francisci Morales, Gianmarco
□ Conrad, Stefan	Deng, Hongbo
☐ Cook, Diane J.	Deng, Kun
☐ Cordeiro, Robson L.F.	Devanur, Nishchal
☐ Crowcroft, Jon	Dhillon, Inderjit S.
☐ Cule, Boris	□ Di Lorenzo, Giusy
. .	□ Ding, Chris
<u>d</u>	☐ Ding, Wei
de Vries, Timothy	Do, Trong Dinh Thac
D	Domeniconi, Carlotta
D	Dou, Dejing
□ Da Silva, Alzennyr	Drumond, Lucas
☐ Dai, Bo	Du, Jun
Das, Kamalika	







Du, Lan	☐ Färber, Ines
□ Du, Liang	☐ Feelders, Ad
□ Duan, Rong	☐ Feng, Boqin
Duivesteijn, Wouter	☐ Ferri, Cèsar
_	☐ Fraisseix, N.
<u>E</u>	☐ Frélicot, C.
☐ Ebadollahi, Shahram	☐ Freudenthaler, Christoph
☐ Elghazel, Haytham	☐ Froehlich, Jon
☐ Eliassi-Rad, Tina	☐ Fu, Ada Wai-Chee
☐ Elkan, Charles	☐ Fu, Hongping
☐ Epps, Julien	
_	<u>G</u>
<u>F</u>	☐ Gantner, Zeno
☐ Faloutsos, Christos	☐ Gao, Huiji
☐ Fan, Ju	☐ Gao, Jing
☐ Fan, Wei	☐ Gao, Xuedong
☐ Fang, Yue	☐ Garboni, Calin







☐ Giannakopoulos, George	Han, Xin
☐ Gieseke, Fabian	Han, Yue
☐ Girgin, Sertan	Haverkamp, Donna S.
☐ Goethals, Bart	Hayashi, Kohei
☐ Goodrich, Ben	He, Jingrui
☐ Gopalkrishnan, Vivekanand	He, Qi
☐ Gullo, Francesco	He, Qing
☐ Günnemann, Stephan	He, Xianmang
☐ Guo, Fan	Hebrail, Georges
☐ Guo, Li	Hernández-Orallo, José
☐ Guo, Yuanyuan	Hoens, T. Ryan
☐ Gwadera, Robert	Hoi, Steven C.H.
	Honavar, Vasant
<u>H</u>	Horne, James H.
☐ Hämäläinen, Wilhelmiina	Hou, Rui
☐ Han, Eui-Hong	Houle, Michael E.
☐ Han, Jiawei	







☐ Hu, Baogang	. I
☐ Hu, Jianying	<u> </u>
☐ Hu, Zhen	Jain, Ramesh
☐ Huan, Jun	Jaroszewicz, Szymon
☐ Huang, Heng	Jensen, David
☐ Huang, Jianbin	Jiang, Peng
☐ Huang, Yalou	Jiang, Xiaohong
☐ Hughes, Ellen K.	Jiang, Yansheng
☐ Hung, Chih-Chieh	Jin, Huidong
☐ Hutchinson, Rebecca A.	Jin, Xiaoming
	K
□ Ikeda, Kazushi	 □ Kabbur, Santosh
☐ Ioffe, Sergey	Kabir, Md. Enamul
☐ Ishigaki, Tsukasa	□ Kalousis, Alexandros
_ isingani, roundou	☐ Kamiran, Faisal
	□ Kamiya, Yuki







☐ Kang, U.	Krzywicki, Alfred
□ Karypis, George	Kukar, Matjaž
☐ Kashima, Hisashi	☐ Kunieda, Kazuo
□ Kaski, Petteri	☐ Kutsuna, Takuro
☐ Kato, Daishi	<u>-</u>
☐ Keogh, Eamonn	<u>L</u>
☐ Khan, Latifur	☐ Lai, Jian-Huang
☐ Kim, Gunhee	☐ Lam, Wai
☐ Kim, Heeyoung	□ Lane, Terran
☐ Kim, Sangkyum	□ Lathia, Neal
☐ Kim, Sungchul	□ Laurent, Anne
☐ Kim, Yang Sok	□ Le Capitaine, H.
☐ Kitts, Brendan	□ Lechevallier, Yves
☐ Knobbe, Arno	☐ Leskovec, Jure
☐ Kong, Xiangnan	Leung, Cane Wing-Ki
König, Arnd Christian	☐ Li, Guangxia
	Li, Juanzi







☐ Li, Ping ☐ Li, Tao ☐ Li, Wei □ Li, Yujia ☐ Li, Zhongmou ☐ Liang, Wenxin ☐ Lim, Ee-Peng ☐ Lin, Qihang ☐ Ling, Charles X. ☐ Liu, Bo ☐ Liu, Huan ☐ Liu, Jie ☐ Liu, Jungiang ☐ Liu, Ning ☐ Liu, Qi

- ☐ Liu, Tantan
- ☐ Liu, Yanchi
- Long, Mingsheng
- Lowd, Daniel
- ☐ Lu, Ming-Yu
- ☐ Lu, Zhengdong
- ☐ Lucchese, Claudio
- ☐ Luo, Dijun
- ☐ Luo, Ping

M

- Ma, Guang-Qin
- Ma, Tengfei
- Mahidadia, Ashesh
- Mahnic, Viljan
- Majumder, Anirban
- Malher, Sébastien



☐ Liu, Sen





■ Mary, Jeremie	N
■ Masseglia, Florent	
■ Masud, Mohammad M.	☐ Najjar, Walid
■ Matsuzaki, Takuya	■ Navas, Mario
☐ McGlohon, Mary	□ Nettleton, Eric
☐ McLachlan, Geoff	□ Nguyen, Nam
☐ Mei, Qiaozhu	□ Nicol, Olivier
☐ Menon, Aditya Krishna	□ Niennattrakul, Vit
☐ Mestre, Julian	□ Niu, Gang
☐ Miettinen, Pauli	□ Niu, Xiaoda
■ Mojsilovic, Aleksandra	Niu, Zhendong
☐ Moruz, Gabriel	Nørvåg, Kjetil
☐ Motomura, Yoichi	■ Noto, Keith
☐ Mueen, Abdullah	
☐ Müller, Emmanuel	<u>U</u>
☐ Muthukrishnan, Pradeep	□ Ojala, Markus
- Mathamonian, Fradoop	□ Ordonez Carlos







☐ Oza, Nikunj C.	Q
P	☐ Qi, Yanjun
☐ Pagh, Rasmus	☐ Qin, Han
☐ Palpanas, Themis	☐ Qiu, Yang
☐ Park, Sunho	Quercia, Daniele
☐ Pathak, Nishith	Quèré, R.
□ Pechenizkiy, Mykola	D
☐ Pei, Jian	<u>R</u>
☐ Peng, Wen-Chih	Radev, Dragomir
☐ Plis, Sergey M.	Raeder, Troy
□ Ponti, Giovanni	Ramírez-Quintana, María José
☐ Powter, Amanda	Ramos, Fabio
□ Poženel, Marko	Rao, Jinghai
☐ Prakash, B. Aditya	Rashidi, Parisa
☐ Preux, Philippe	Ratanamahatana, Chotirat Ann
Troux, rrimppe	Rattigan, Matthew J.H.







- ☐ Rendle, Steffen
- □ Rzepakowski, Piotr

S

- Sarne, David
- ☐ Sart, Doruk
- Savas, Berkant
- ☐ Schlüter, Tim
- Schmidt-Thieme, Lars
- □ Scott, Stephen
- ☐ Sebag, Michèle
- Seidl, Thomas
- ☐ Seufert, Stephan
- ☐ Shafiei, Mahdi
- ☐ Shaham, Eran
- ☐ Shan, Hanhuai
- ☐ Shekhar, Shashi

- ☐ Shen, Dou
- ☐ Shen, Xukun
- ☐ Shen, Yidong
- ☐ Shen, Zhiyong
- ☐ Shi, Baile
- ☐ Shi, Xiaoxiao
- ☐ Shi, Zhongzhi
- Shibata, Tomohiro
- ☐ Shieh, Jin
- Shrivastava, Nisheeth
- Silvescu, Adrian
- ☐ Sim, Kelvin
- Sindhwani, Vikas
- Slonim, Donna
- Song, Yangqiu
- ☐ Sow, Daby







☐ Srivastava, Ashok N.	☐ Tao, Dacheng
□ Srivastava, Jaideep	□ Tatti, Nikolaj
☐ Sun, Heli	Telang, Aditya
□ Sun, Jimeng	□ Termier, Alexandre
☐ Sun, Xu	☐ Thiel, Kilian
■ Svonava, Daniel	Thuraisingham, Bhavani
-	☐ Tian, Jilei
<u>l</u>	☐ Ting, Kai Ming
☐ Tagarelli, Andrea	☐ Tischer, Peter
☐ Takenaka, Takeshi	☐ Toivonen, Hannu
☐ Takenouchi, Takashi	Tong, Hanghang
☐ Tan, Jianlong	Traina Jr., Caetano
☐ Tang, Jie	☐ Traina, Agma J.M.
☐ Tang, Lei	☐ Tsang, Ivor
☐ Tang, Liang	☐ Tsourakakis, Charalampos
☐ Tang, Lu-An	



☐ Tang, Wei





Wang, Chang-Dong
Wang, Chun-Li
Wang, Dingding
Wang, Fan
Wang, Fei
■ Wang, Hao
■ Wang, Hua
Wang, Jianmin
Wang, Jianyong
■ Wang, Ke
Wang, Peng
Wang, Wei
■ Wang, Xiang
■ Wang, Xufei
Wang, Yanqi
Wang, Zheng







☐ Wei, Liang	☐ Xiao, Yanshan
Weikum, Gerhard	☐ Xie, Yan
□ Wells, Jonathan R.	☐ Xiong, Hui
Weng, Jianshu	Xiong, Yuhong
■ Wobcke, Wayne	Xu, Yabo
Wong, Raymond Chi-Wing	
■ Wong, Tak-Lam	<u>Y</u>
■ Wong, Weng-Keen	Yamada, Keiji
■ Woznica, Adam	Yan, Jun
☐ Wu, Jun	Yan, Shuicheng
□ Wu, Junjie	Yang, Jaewon
■ Wu, Xindong	Yang, Qing
□ Wu, Yao	Yang, Qingyan
	Yang, Shengwen
X	Yang, Zi
☐ Xia, Chaolun	☐ Yildirim, Hilmi
☐ Xiao, Jing	☐ Yin, Jie







	Yu, Hwanjo	Zhang, Min-Ling
	Yu, Jeffrey Xu	Zhang, Peng
	Yu, Jun	Zhang, Xianchao
	Yu, Kai	Zhang, Xiangliang
	Yu, Kui	Zhang, Yanchun
	Yu, Lei	Zhang, Yi
	Yu, Philip S.	Zhang, Yuan
	Yu, Xiao	Zhao, Peixiang
	Yuan, Yifei	Zheng, Li
_		Zheng, Yaling
<u>Z</u>		Zheng, Zeyu
	Zhang, Changshui	Zhou, Aoying
	Zhang, Chongsheng	Zhou, Fang
	Zhang, Chunxia	Zhou, Hang
	Zhang, Harry	Zhou, Lizhu
	Zhang, Li	Zhou, Tianyi
	Zhang, Ming	







- ☐ Zhou, Yang
- ☐ Zhou, Zhi-Hua
- ☐ Zhu, Huisheng
- ☐ Zhu, Jun-Yong
- ☐ Zhu, Qiang
- ☐ Zhu, Ruixin
- ☐ Zhu, Xingquan
- ☐ Zhuang, Fuzhen
- Zou, Jia







Abello, James Detecting Novel Discrepancies in Communication Networks Aggarwal, Charu Addressing Concept-Evolution in Concept-Drifting Data Streams Agrawal, Gagan Stratified Sampling for Data Mining on the Deep Web Akoglu, Leman Patterns on the Connected Components of Terabyte-Scale Graphs Al Hasan, Mohammad





SONNET: Efficient Approximate Nearest Neighbor Using Multi-core



Alamgir, Morteza Multi-agent Random Walks for Local Clustering on Graphs Albrecht, David Accelerating Radius-Margin Parameter Selection for SVMs Using Geometric **Bounds** Au, Dyng Attribution of Conversion Events to Multi-channel Media Au, Tom S. Spatiotemporal Event Detection in Mobility Network Aung, Zeyar





Discovering Correlated Subspace Clusters in 3D Continuous-Valued Data



Aussem, Alex

☐ Feature Selection for Unsupervised Learning Using Random Cluster Ensembles

Bai, Bing

Learning Preferences with Millions of Parameters by Enforcing Sparsity

Bain, Michael

■ Learning Collaborative Filtering and Its Application to People to People Recommendation in Social Networks

Baladandayuthapani, Veerabhadran

 On the Computation of Stochastic Search Variable Selection in Linear Regression with UDFs







Balakrishnan, Suhrid

Two of a Kind or the Ratings Game? Adaptive Pairwise Preferences and Latent Factor Models

Banerjee, Arindam

- □ A Generalized Linear Threshold Model for Multiple Cascades
- Generalized Probabilistic Matrix Factorizations for Collaborative Filtering

Bao, Tengfei

An Unsupervised Approach to Modeling Personalized Contexts of Mobile Users

Baraglia, Ranieri

Document Similarity Self-Join with MapReduce

Bedathur, Srikanta

Bonsai: Growing Interesting Small Trees







Bella, Antonio Quantification via Probability Estimators Ben-Moshe, Boaz Co-clustering of Lagged Data Berthold, Michael R. Node Similarities from Spreading Activation Bhaduri, Kanishka Fast and Flexible Multivariate Time Series Subsequence Search Bhatnagar, Raj





Algorithm for Discovering Low-Variance 3-Clusters from Real-Valued Datasets



Boden, Brigitte

 Subspace Clustering Meets Dense Subgraph Mining: A Synthesis of Two Paradigms

Bogorny, Vania

Spatial and Spatio-temporal Data Mining

Brodley, Carly

Anomaly Detection Using an Ensemble of Feature Models

Bucak, Serhat S.

One-Class Matrix Completion with Low-Density Factorizations

Buntine, Wray Lindsay

 Sequential Latent Dirichlet Allocation: Discover Underlying Topic Structures within a Document







Burdick, Brian Attribution of Conversion Events to Multi-channel Media Cai, Xiongcai Learning Collaborative Filtering and Its Application to People to People Recommendation in Social Networks Calabrese, Francesco Recommending Social Events from Mobile Phone Location Data Calders, Toon Approximation of Frequentness Probability of Itemsets in Uncertain Data Discrimination Aware Decision Tree Learning Calhoun, Vince D.





Permutations as Angular Data: Efficient Inference in Factorial Spaces



Camerra, Alessandro

☐ iSAX 2.0: Indexing and Mining One Billion Time Series

Campagna, Andrea

On Finding Frequent Patterns in Event Sequences

Cao, Happia

■ An Unsupervised Approach to Modeling Personalized Contexts of Mobile Users

Cao, Longbing

- ☐ Exploiting Local Data Uncertainty to Boost Global Outlier Detection
- SMILE: A Similarity-Based Approach for Multiple Instance Learning

Capra, Licia

Mining Public Transport Usage for Personalised Intelligent Transport Systems







Caragea, Cornelia	
□ Abstraction Augmented Markov Models	
Caragea, Doina	
□ Abstraction Augmented Markov Models	
Carbonell, Jaime	
□ Rare Category Characterization	
☐ Learning Preferences with Millions of Parameters by Enforcing Sparsity	
Cebron, Nicolas	
☐ Active Improvement of Hierarchical Object Features under Budget Constraints	
Chakraborty, Abhirup	
□ SONNET: Efficient Approximate Nearest Neighbor Using Multi-core	







Chakravarthy, Sharma A Graph-Based Approach for Multi-folder Email Classification Chan, Shing-Kit Pseudo Conditional Random Fields: Joint Training Approach to Segmenting and Labeling Sequence Data Chang, Kuiyu Micro-blogging Sentiment Detection by Collaborative Online Learning Chau, Duen Horng On the Vulnerability of Large Graphs Chawla, Nitesh V.





Consequences of Variability in Classifier Performance Estimates



Chawla, Sanjay Finding Local Anomalies in Very High Dimensional Space Chen, Bo Location and Scatter Matching for Dataset Shift in Text Mining Chen, Enhong An Unsupervised Approach to Modeling Personalized Contexts of Mobile Users Chen, Qing Addressing Concept-Evolution in Concept-Drifting Data Streams Chen, Wei Scalable Influence Maximization in Social Networks under the Linear Threshold



Model





Chen, Xi
☐ Learning Preferences with Millions of Parameters by Enforcing Sparsity
Chen, Yiran
☐ MoodCast: Emotion Prediction via Dynamic Continuous Factor Graph Model
Chen, Zheng
□ A Novel Contrast Co-learning Framework for Generating High Quality Training Data
Cheng, Hong
☐ Clustering Large Attributed Graphs: An Efficient Incremental Approach
Cheng, Wei
☐ Transfer Learning via Cluster Correspondence Inference







Chiky, Raja

 CLUSMASTER: A Clustering Approach for Sampling Data Streams in Sensor Networks

Chipman, Hugh

Mixed-Membership Stochastic Block-Models for Transactional Networks

Choi, Seungjin

Bayesian Aggregation of Binary Classifiers

Chopra, Sumit

□ Two of a Kind or the Ratings Game? Adaptive Pairwise Preferences and Latent Factor Models







Compton, Paul

■ Learning Collaborative Filtering and Its Application to People to People Recommendation in Social Networks

Conrad, Stefan

□ An Approach for Automatic Sleep Stage Scoring and Apnea-Hypopnea Detection

Cook, Diane J.

Mining Sensor Streams for Discovering Human Activity Patterns over Time

Cordeiro, Robson L.F.

QMAS: Querying, Mining and Summarization of Multi-modal Databases

Crowcroft, Jon

Recommending Social Events from Mobile Phone Location Data







Cule, Boris Mining Closed Strict Episodes Da Silva, Alzennyr CLUSMASTER: A Clustering Approach for Sampling Data Streams in Sensor **Networks** Dai, Bo **Bayesian Maximum Margin Clustering** Das, Kamalika Block-GP: Scalable Gaussian Process Regression for Multimodal Data Datta, Samik Viral Marketing for Multiple Products







Davis, Jesse □ Learning Markov Network Structure with Decision Trees De Francisci Morales, Gianmarco □ Document Similarity Self-Join with MapReduce de Vries, Timothy □ Finding Local Anomalies in Very High Dimensional Space



gSkeletonClu: Density-Based Network Clustering via Structure-Connected Tree
 Division or Agglomeration







Deng, Kun Active Learning from Multiple Noisy Labelers with Varied Costs **Devanur, Nishchal Detecting Novel Discrepancies in Communication Networks** Dhillon, Inderjit S. Supervised Link Prediction Using Multiple Sources Di Lorenzo, Giusy Recommending Social Events from Mobile Phone Location Data Ding, Chris Towards Structural Sparsity: An Explicit 12/10 Approach Weighted Feature Subset Non-negative Matrix Factorization and Its Applications to Document Understanding







☐ Hierarchical Ensemble Clustering
Ding, Wei
□ Causal Discovery from Streaming Features
Do, Trong Dinh Thac
□ PGLCM: Efficient Parallel Mining of Closed Frequent Gradual Itemsets
Domeniconi, Carlotta
☐ Enhancing Single-Objective Projective Clustering Ensembles
Enhancing Single-Objective Projective Clustering EnsemblesDou, Dejing
Dou, Dejing







Du, Jun Active Learning with Human-Like Noisy Oracle Du, Lan Sequential Latent Dirichlet Allocation: Discover Underlying Topic Structures within a Document Du, Liang Interval-valued Matrix Factorization with Applications Duan, Rong Spatiotemporal Event Detection in Mobility Network Duivesteijn, Wouter Subgroup Discovery Meets Bayesian Networks—An Exceptional Model Mining Approach







Ebadollahi, Shahram

□ A System for Mining Temporal Physiological Data Streams for Advanced Prognostic Decision Support

Elghazel, Haytham

☐ Feature Selection for Unsupervised Learning Using Random Cluster Ensembles

Eliassi-Rad, Tina

- Detecting Novel Discrepancies in Communication Networks
- On the Vulnerability of Large Graphs

Elkan, Charles

A Log-Linear Model with Latent Features for Dyadic Prediction







Epps, Julien

 minCEntropy: A Novel Information Theoretic Approach for the Generation of Alternative Clusterings

Faloutsos, Christos

- Mining Billion-node Graphs: Patterns, Generators and Tools
- QMAS: Querying, Mining and Summarization of Multi-modal Databases
- Patterns on the Connected Components of Terabyte-Scale Graphs
- On the Vulnerability of Large Graphs

Fan, Ju

 Personalizing Web Page Recommendation via Collaborative Filtering and Topic-Aware Markov Model

Fan, Wei

■ Efficient Semi-supervised Spectral Co-clustering with Constraints







	Transfer Learning on Heterogenous Feature Spaces via Spectral Transformation
Fa	ing, Yue
	Financial Forecasting with Gompertz Multiple Kernel Learning
Fä	irber, Ines
	Subspace Clustering Meets Dense Subgraph Mining: A Synthesis of Two Paradigms
	Discovering Multiple Clustering Solutions: Grouping Objects in Different Views of the Data
Fe	elders, Ad
	Subgroup Discovery Meets Bayesian Networks—An Exceptional Model Mining Approach
	Monotone Relabeling in Ordinal Classification







Feng, Boqin

gSkeletonClu: Density-Based Network Clustering via Structure-Connected Tree
 Division or Agglomeration

Ferri, Cèsar

Quantification via Probability Estimators

Fraisseix, N.

 On Normalizing Fuzzy Coincidence Matrices to Compare Fuzzy and/or Possibilistic Partitions with the Rand Index

Frélicot, C.

 On Normalizing Fuzzy Coincidence Matrices to Compare Fuzzy and/or Possibilistic Partitions with the Rand Index

Freudenthaler, Christoph

Learning Attribute-to-Feature Mappings for Cold-Start Recommendations







Froehlich, Jon Mining Public Transport Usage for Personalised Intelligent Transport Systems Fu, Ada Wai-Chee **Anonymizing Temporal Data** Probabilistic Inference Protection on Anonymized Data Fu, Hongping An Approach Based on Tree Kernels for Opinion Mining of Online Product Reviews Gantner, Zeno Learning Attribute-to-Feature Mappings for Cold-Start Recommendations Gao, Huiji





Discovering Overlapping Groups in Social Media



Gao, Jing Addressing Concept-Evolution in Concept-Drifting Data Streams Gao, Xuedong Understanding of Internal Clustering Validation Measures Garboni, Calin Approximation of Frequentness Probability of Itemsets in Uncertain Data Giannakopoulos, George The Effect of History on Modeling Systems' Performance: The Problem of the Demanding Lord Gieseke, Fabian Resilient K-d Trees: K-Means in Space Revisited







Girgin, Sertan Advertising Campaigns Management: Should We Be Greedy? Goethals, Bart Approximation of Frequentness Probability of Itemsets in Uncertain Data Goodrich, Ben Accelerating Radius-Margin Parameter Selection for SVMs Using Geometric **Bounds** Gopalkrishnan, Vivekanand Discovering Correlated Subspace Clusters in 3D Continuous-Valued Data **Gullo, Francesco**





Enhancing Single-Objective Projective Clustering Ensembles



	Minimizing the Variance of Cluster Mixture Models for Clustering Uncertain Objects
Gi	innemann, Stephan
	Subspace Clustering Meets Dense Subgraph Mining: A Synthesis of Two Paradigms
	Discovering Multiple Clustering Solutions: Grouping Objects in Different Views of the Data
Gı	uo, Fan
	QMAS: Querying, Mining and Summarization of Multi-modal Databases
Gı	uo, Li
	Classifier and Cluster Ensembles for Mining Concept Drifting Data Streams



Guo, Yuanyuan



An Extensive Empirical Study on Semi-supervised Learning



Gwadera, Robert

Multi-stream Join Answering for Mining Significant Cross-Stream Correlations

Hämäläinen, Wilhelmiina

Efficient Discovery of the Top-K Optimal Dependency Rules with Fisher's Exact Test of Significance

Han, Eui-Hong

Content-Based Methods for Predicting Web-Site Demographic Attributes

Han, Jiawei

- gSkeletonClu: Density-Based Network Clustering via Structure-Connected Tree Division or Agglomeration
- Addressing Concept-Evolution in Concept-Drifting Data Streams
- ☐ Tru-Alarm: Trustworthiness Analysis of Sensor Networks in Cyber-Physical Systems







Han, Xin
☐ Graph-Based Semi-supervised Learning with Adaptive Similarity Estimation
Han, Yue
□ A Variance Reduction Framework for Stable Feature Selection
Haverkamp, Donna S.
☐ QMAS: Querying, Mining and Summarization of Multi-modal Databases
Hayashi, Kohei
 Exponential Family Tensor Factorization for Missing-Values Prediction and Anomaly Detection
He, Jingrui
□ Rare Category Characterization







He, Qi What Do People Want in Microblogs? Measuring Interestingness of Hashtags in Twitter He, Qing D-LDA: A Topic Modeling Approach without Constraint Generation for Semi-defined Classification He, Xianmang Efficient Episode Mining with Minimal and Non-overlapping Occurrences Hebrail, Georges CLUSMASTER: A Clustering Approach for Sampling Data Streams in Sensor **Networks** Hernández-Orallo, José





Quantification via Probability Estimators



Hoens, T. Ryan Consequences of Variability in Classifier Performance Estimates Hoi, Steven C.H. Micro-blogging Sentiment Detection by Collaborative Online Learning Honavar, Vasant Abstraction Augmented Markov Models Horne, James H. QMAS: Querying, Mining and Summarization of Multi-modal Databases Hou, Rui





Frequent Instruction Sequential Pattern Mining in Hardware Sample Data



Houle, Michael E. I Finding Local Anomalies in Very High Dimensional Space Hu, Baogang Bayesian Maximum Margin Clustering Hu, Jianying One-Class Matrix Completion with Low-Density Factorizations A System for Mining Temporal Physiological Data Streams for Advanced Prognostic Decision Support Hu, Zhen

Algorithm for Discovering Low-Variance 3-Clusters from Real-Valued Datasets







Huan, Jun □ Knowledge Discovery in Academic Drug Discovery Programs: Opportunities and Challenges Huang, Heng

☐ Towards Structural Sparsity: An Explicit 12/10 Approach

Huang, Jianbin

gSkeletonClu: Density-Based Network Clustering via Structure-Connected Tree
 Division or Agglomeration

Huang, Yalou

 Training Conditional Random Fields Using Transfer Learning for Gesture Recognition

Hughes, Ellen K.

QMAS: Querying, Mining and Summarization of Multi-modal Databases







Hung, Chih-Chieh

 Tru-Alarm: Trustworthiness Analysis of Sensor Networks in Cyber-Physical Systems

Hutchinson, Rebecca A.

Modeling Experts and Novices in Citizen Science Data for Species Distribution Modeling

Ikeda, Kazushi

■ Exponential Family Tensor Factorization for Missing-Values Prediction and Anomaly Detection

Ioffe, Sergey

Improved Consistent Sampling, Weighted Minhash and L1 Sketching







Ishigaki, Tsukasa Category Mining by Heterogeneous Data Fusion Using PdLSI Model in a Retail Service Jain, Ramesh Micro-blogging Sentiment Detection by Collaborative Online Learning Jaroszewicz, Szymon **Decision Trees for Uplift Modeling** Jensen, David Leveraging D-Separation for Relational Data Sets Jiang, Peng



Reviews



An Approach Based on Tree Kernels for Opinion Mining of Online Product



Jiang, Xiaohong		
☐ Efficient Probabilistic Latent Semantic Analysis with Sparsity Control		
Jiang, Yansheng		
☐ Graph-Based Semi-supervised Learning with Adaptive Similarity Estimation		
Jin, Huidong		
 Sequential Latent Dirichlet Allocation: Discover Underlying Topic Structures within a Document 		
Jin, Xiaoming		
☐ Transfer Learning via Cluster Correspondence Inference		
Kabbur, Santosh		
☐ Content-Based Methods for Predicting Web-Site Demographic Attributes		







Kabir, Md. Enamul A Pairwise-Systematic Microaggregation for Statistical Disclosure Control Kalousis, Alexandros Adaptive Distances on Sets of Vectors Kamiran, Faisal Discrimination Aware Decision Tree Learning Kamiya, Yuki Exponential Family Tensor Factorization for Missing-Values Prediction and Anomaly Detection Kang, U. Patterns on the Connected Components of Terabyte-Scale Graphs







Karypis, George

Content-Based Methods for Predicting Web-Site Demographic Attributes

Kashima, Hisashi

 Averaged Stochastic Gradient Descent with Feedback: An Accurate, Robust, and Fast Training Method

Kaski, Petteri

☐ Testing the Significance of Patterns in Data with Cluster Structure

Kato, Daishi

 Exponential Family Tensor Factorization for Missing-Values Prediction and Anomaly Detection

Keogh, Eamonn

□ iSAX 2.0: Indexing and Mining One Billion Time Series







Polishing the Right Apple: Anytime Classification Also Benefits Data Streams with Constant Arrival Times Mother Fugger: Mining Historical Manuscripts with Local Color Patches Data Editing Techniques to Allow the Application of Distance-Based Outlier **Detection to Streams** Accelerating Dynamic Time Warping Subsequence Search with GPUs and FPGAs How to Do Good Data Mining Research and Get it Published in Top Venues Khan, Latifur Addressing Concept-Evolution in Concept-Drifting Data Streams Kim, Gunhee QMAS: Querying, Mining and Summarization of Multi-modal Databases Kim, Heeyoung Spatiotemporal Event Detection in Mobility Network







Kim, Sangkyum

 Tru-Alarm: Trustworthiness Analysis of Sensor Networks in Cyber-Physical Systems

Kim, Sungchul

□ Passive Sampling for Regression

Kim, Yang Sok

Learning Collaborative Filtering and Its Application to People to People Recommendation in Social Networks

Kitts, Brendan

Attribution of Conversion Events to Multi-channel Media







Knobbe, Arno

 Subgroup Discovery Meets Bayesian Networks—An Exceptional Model Mining Approach

Kong, Xiangnan

■ Multi-label Feature Selection for Graph Classification

König, Arnd Christian

☐ Learning a Bi-Stochastic Data Similarity Matrix

Krzywicki, Alfred

■ Learning Collaborative Filtering and Its Application to People to People Recommendation in Social Networks

Kukar, Matjaž

□ Separation of Interleaved Web Sessions with Heuristic Search







Kunieda, Kazuo

 Exponential Family Tensor Factorization for Missing-Values Prediction and Anomaly Detection

Kutsuna, Takuro

A Binary Decision Diagram-Based One-Class Classifier

Lai, Jian-Huang

□ A Conscience On-line Learning Approach for Kernel-Based Clustering

Lam, Wai

- Pseudo Conditional Random Fields: Joint Training Approach to Segmenting and Labeling Sequence Data
- Location and Scatter Matching for Dataset Shift in Text Mining







Lane, Terran Permutations as Angular Data: Efficient Inference in Factorial Spaces Lathia, Neal Mining Public Transport Usage for Personalised Intelligent Transport Systems Recommending Social Events from Mobile Phone Location Data Laurent, Anne PGLCM: Efficient Parallel Mining of Closed Frequent Gradual Itemsets Le Capitaine, H. On Normalizing Fuzzy Coincidence Matrices to Compare Fuzzy and/or Possibilistic Partitions with the Rand Index Lechevallier, Yves





ABS: The Anti Bouncing Model for Usage Data Streams



Leskovec, Jure Modeling Information Diffusion in Implicit Networks Leung, Cane Wing-Ki What Do People Want in Microblogs? Measuring Interestingness of Hashtags in *Twitte*r Li, Guangxia Micro-blogging Sentiment Detection by Collaborative Online Learning Li, Juanzi Term Filtering with Bounded Error Li, Ping





Learning a Bi-Stochastic Data Similarity Matrix

Compressed Nonnegative Sparse Coding



Li, Tao LogTree: A Framework for Generating System Events from Raw Textual Logs Weighted Feature Subset Non-negative Matrix Factorization and Its Applications to Document Understanding Hierarchical Ensemble Clustering Li, Wei Term Filtering with Bounded Error Li, Yujia Efficient Episode Mining with Minimal and Non-overlapping Occurrences Li, Zhongmou Detecting Blackhole and Volcano Patterns in Directed Networks





Understanding of Internal Clustering Validation Measures



Liang, Wenxin Graph-Based Semi-supervised Learning with Adaptive Similarity Estimation Lim, Ee-Peng What Do People Want in Microblogs? Measuring Interestingness of Hashtags in *Twitte*r Lin, Qihang Learning Preferences with Millions of Parameters by Enforcing Sparsity Ling, Charles X. Active Learning with Human-Like Noisy Oracle Liu, Bo Exploiting Local Data Uncertainty to Boost Global Outlier Detection





SMILE: A Similarity-Based Approach for Multiple Instance Learning



Liu, Huan Discovering Overlapping Groups in Social Media Liu, Jie Training Conditional Random Fields Using Transfer Learning for Gesture Recognition Liu, Junqiang Enforcing Vocabulary k-Anonymity by Semantic Similarity Based Clustering Liu, Ning A Novel Contrast Co-learning Framework for Generating High Quality Training Data Liu, Qi





Transfer Learning on Heterogenous Feature Spaces via Spectral Transformation



Liu, Sen Efficient Probabilistic Latent Semantic Analysis with Sparsity Control Liu, Tantan Stratified Sampling for Data Mining on the Deep Web Liu, Yanchi Detecting Blackhole and Volcano Patterns in Directed Networks Understanding of Internal Clustering Validation Measures Long, Mingsheng Transfer Learning via Cluster Correspondence Inference Lowd, Daniel





Learning Markov Network Structure with Decision Trees



Lu, Ming-Yu Collaborative Learning between Visual Content and Hidden Semantic for Image Retrieval Lu, Zhengdong Supervised Link Prediction Using Multiple Sources

Lucchese, Claudio

Document Similarity Self-Join with MapReduce

Luo, Dijun

☐ Towards Structural Sparsity: An Explicit 12/10 Approach

Luo, Ping

□ D-LDA: A Topic Modeling Approach without Constraint Generation for Semi-defined Classification







□ Topic Modeling Ensembles	
Ma, Guang-Qin	
□ Spatiotemporal Event Detection in Mobility Network	
Ma, Tengfei	
☐ Multi-document Summarization Using Minimum Distortion	
Mahidadia, Ashesh	
□ Learning Collaborative Filtering and Its Application to People to People Recommendation in Social Networks	
Mahnic, Viljan	
□ Separation of Interleaved Web Sessions with Heuristic Search	
Majumder, Anirban	
□ Viral Marketing for Multiple Products	







Malher, Sébastien Network Simplification with Minimal Loss of Connectivity Mary, Jeremie Advertising Campaigns Management: Should We Be Greedy? Masseglia, Florent ABS: The Anti Bouncing Model for Usage Data Streams Masud, Mohammad M. Addressing Concept-Evolution in Concept-Drifting Data Streams Matsuzaki, Takuya Averaged Stochastic Gradient Descent with Feedback: An Accurate, Robust,



and Fast Training Method





McGlohon, Mary □ Patterns on the Connected Components of Terabyte-Scale Graphs McLachlan, Geoff □ Assessing the Significance of Groups in High-Dimensional Data Mei, Qiaozhu □ Edge Weight Regularization over Multiple Graphs for Similarity Learning Menon, Aditya Krishna □ A Log-Linear Model with Latent Features for Dyadic Prediction Mestre, Julian





Bonsai: Growing Interesting Small Trees



Miettinen, Pauli Sparse Boolean Matrix Factorizations Mojsilovic, Aleksandra One-Class Matrix Completion with Low-Density Factorizations Moruz, Gabriel Resilient K-d Trees: K-Means in Space Revisited Motomura, Yoichi Category Mining by Heterogeneous Data Fusion Using PdLSI Model in a Retail Service Mueen, Abdullah Accelerating Dynamic Time Warping Subsequence Search with GPUs



and FPGAs





Müller, Emmanuel

 Discovering Multiple Clustering Solutions: Grouping Objects in Different Views of the Data

Muthukrishnan, Pradeep

Edge Weight Regularization over Multiple Graphs for Similarity Learning

Najjar, Walid

Accelerating Dynamic Time Warping Subsequence Search with GPUs and FPGAs

Navas, Mario

 On the Computation of Stochastic Search Variable Selection in Linear Regression with UDFs

Nettleton, Eric

Improving Kernel Methods through Complex Data Mapping







Nguyen, Nam □ A New SVM Approach to Multi-instance Multi-label Learning Nicol, Olivier □ Advertising Campaigns Management: Should We Be Greedy? Niennattrakul, Vit □ Data Editing Techniques to Allow the Application of Distance-Based Outlier Detection to Streams □ Accelerating Dynamic Time Warping Subsequence Search with GPUs and FPGAs

Niu, Gang

Bayesian Maximum Margin Clustering







Niu, Xiaoda An Extensive Empirical Study on Semi-supervised Learning Niu, Zhendong An Approach Based on Tree Kernels for Opinion Mining of Online Product Reviews Nørvåg, Kjetil K-AP: Generating Specified K Clusters by Efficient Affinity Propagation Noto, Keith Anomaly Detection Using an Ensemble of Feature Models Ojala, Markus Assessing Data Mining Results on Matrices with Randomization







Ordonez, Carlos

 On the Computation of Stochastic Search Variable Selection in Linear Regression with UDFs

Oza, Nikunj C.

☐ Fast and Flexible Multivariate Time Series Subsequence Search

Pagh, Rasmus

On Finding Frequent Patterns in Event Sequences

Palpanas, Themis

- ☐ iSAX 2.0: Indexing and Mining One Billion Time Series
- □ The Effect of History on Modeling Systems' Performance: The Problem of the Demanding Lord







Park, Sunho **Bayesian Aggregation of Binary Classifiers** Pathak, Nishith A Generalized Linear Threshold Model for Multiple Cascades Pechenizkiy, Mykola Discrimination Aware Decision Tree Learning Pei, Jian Probabilistic Inference Protection on Anonymized Data Peng, Wen-Chih Tru-Alarm: Trustworthiness Analysis of Sensor Networks in Cyber-Physical **Systems**







Plis, Sergey M. Permutations as Angular Data: Efficient Inference in Factorial Spaces Ponti, Giovanni Minimizing the Variance of Cluster Mixture Models for Clustering Uncertain **Objects** Powter, Amanda Attribution of Conversion Events to Multi-channel Media Poženel, Marko Separation of Interleaved Web Sessions with Heuristic Search Prakash, B. Aditya





On the Vulnerability of Large Graphs



Preux, Philippe Advertising Campaigns Management: Should We Be Greedy? Qi, Yanjun Learning Preferences with Millions of Parameters by Enforcing Sparsity Qin, Han Financial Forecasting with Gompertz Multiple Kernel Learning Qiu, Yang Constraint Based Dimension Correlation and Distance Divergence for Clustering **High-Dimensional Data** Quercia, Daniele





Recommending Social Events from Mobile Phone Location Data



Quèré, R. On Normalizing Fuzzy Coincidence Matrices to Compare Fuzzy and/or Possibilistic Partitions with the Rand Index Radev, Dragomir Edge Weight Regularization over Multiple Graphs for Similarity Learning Raeder, Troy Consequences of Variability in Classifier Performance Estimates Ramírez-Quintana, María José Quantification via Probability Estimators Ramos, Fabio





Improving Kernel Methods through Complex Data Mapping



Rao, Jinghai MoodCast: Emotion Prediction via Dynamic Continuous Factor Graph Model Rashidi, Parisa Mining Sensor Streams for Discovering Human Activity Patterns over Time Ratanamahatana, Chotirat Ann Data Editing Techniques to Allow the Application of Distance-Based Outlier **Detection to Streams** Rattigan, Matthew J.H. Leveraging D-Separation for Relational Data Sets Rendle, Steffen Learning Attribute-to-Feature Mappings for Cold-Start Recommendations **Factorization Machines**







Rzepakowski, Piotr **Decision Trees for Uplift Modeling** Sarne, David Co-clustering of Lagged Data Sart, Doruk Accelerating Dynamic Time Warping Subsequence Search with GPUs and FPGAs Savas, Berkant Supervised Link Prediction Using Multiple Sources Schlüter, Tim An Approach for Automatic Sleep Stage Scoring and Apnea-Hypopnea Detection







Schmidt-Thieme, Lars

Learning Attribute-to-Feature Mappings for Cold-Start Recommendations

Scott, Stephen

□ Active Learning from Multiple Noisy Labelers with Varied Costs

Sebag, Michèle

☐ *K*-AP: Generating Specified *K* Clusters by Efficient Affinity Propagation

Seidl, Thomas

- Subspace Clustering Meets Dense Subgraph Mining: A Synthesis of Two Paradigms
- Discovering Multiple Clustering Solutions: Grouping Objects in Different Views of the Data







Seufert, Stephan Bonsai: Growing Interesting Small Trees Shafiei, Mahdi Mixed-Membership Stochastic Block-Models for Transactional Networks Shaham, Eran Co-clustering of Lagged Data Shan, Hanhuai Generalized Probabilistic Matrix Factorizations for Collaborative Filtering Shekhar, Shashi





Spatial and Spatio-temporal Data Mining



Shen, Dou		
	Transfer Learning via Cluster Correspondence Inference	
Shen, Xukun		
	Topic Modeling Ensembles	
	Interval-valued Matrix Factorization with Applications	
Shen, Yidong		
	Interval-valued Matrix Factorization with Applications	
Shen, Zhiyong		
	D-LDA: A Topic Modeling Approach without Constraint Generation for Semi-defined Classification	
	Topic Modeling Ensembles	
	Interval-valued Matrix Factorization with Applications	







Shi, Baile

Efficient Episode Mining with Minimal and Non-overlapping Occurrences

Shi, Xiaoxiao

- Efficient Semi-supervised Spectral Co-clustering with Constraints
- ☐ Transfer Learning on Heterogenous Feature Spaces via Spectral Transformation

Shi, Zhongzhi

□ D-LDA: A Topic Modeling Approach without Constraint Generation for Semi-defined Classification

Shibata, Tomohiro

 Exponential Family Tensor Factorization for Missing-Values Prediction and Anomaly Detection







Shieh, Jin

- ☐ iSAX 2.0: Indexing and Mining One Billion Time Series
- Polishing the Right Apple: Anytime Classification Also Benefits Data Streams with Constant Arrival Times

Shrivastava, Nisheeth

Viral Marketing for Multiple Products

Silvescu, Adrian

Abstraction Augmented Markov Models

Sim, Kelvin

Discovering Correlated Subspace Clusters in 3D Continuous-Valued Data

Sindhwani, Vikas

One-Class Matrix Completion with Low-Density Factorizations







Slonim, Donna Anomaly Detection Using an Ensemble of Feature Models Song, Yangqiu Homotopy Regularization for Boosting Sow, Daby A System for Mining Temporal Physiological Data Streams for Advanced **Prognostic Decision Support** Srivastava, Ashok N. Fast and Flexible Multivariate Time Series Subsequence Search Block-GP: Scalable Gaussian Process Regression for Multimodal Data Srivastava, Jaideep





A Generalized Linear Threshold Model for Multiple Cascades



Sun, Heli

gSkeletonClu: Density-Based Network Clustering via Structure-Connected Tree
 Division or Agglomeration

Sun, Jimeng

- A System for Mining Temporal Physiological Data Streams for Advanced Prognostic Decision Support
- MoodCast: Emotion Prediction via Dynamic Continuous Factor Graph Model

Sun, Xu

 Averaged Stochastic Gradient Descent with Feedback: An Accurate, Robust, and Fast Training Method

Svonava, Daniel

Visualizing Graphs Using Minimum Spanning Dendrograms







Tagarelli, Andrea

- Enhancing Single-Objective Projective Clustering Ensembles
- Minimizing the Variance of Cluster Mixture Models for Clustering Uncertain Objects

Takenaka, Takeshi

Category Mining by Heterogeneous Data Fusion Using PdLSI Model in a Retail Service

Takenouchi, Takashi

 Exponential Family Tensor Factorization for Missing-Values Prediction and Anomaly Detection

Tan, Jianlong

Classifier and Cluster Ensembles for Mining Concept Drifting Data Streams







Tang, Jie Term Filtering with Bounded Error MoodCast: Emotion Prediction via Dynamic Continuous Factor Graph Model Tang, Lei Discovering Overlapping Groups in Social Media Tang, Liang LogTree: A Framework for Generating System Events from Raw Textual Logs Tang, Lu-An Tru-Alarm: Trustworthiness Analysis of Sensor Networks in Cyber-Physical **Systems** Tang, Wei





Supervised Link Prediction Using Multiple Sources



Tao, Dacheng		
□ NESVM: A Fast Gradient Method for Support Vector Machines		
Tatti, Nikolaj		
☐ Mining Closed Strict Episodes		
Telang, Aditya		
☐ A Graph-Based Approach for Multi-folder Email Classification		
Termier, Alexandre		
□ PGLCM: Efficient Parallel Mining of Closed Frequent Gradual Itemsets		
Thiel, Kilian		
■ Node Similarities from Spreading Activation		







Thuraisingham, Bhavani Addressing Concept-Evolution in Concept-Drifting Data Streams Tian, Jilei An Unsupervised Approach to Modeling Personalized Contexts of Mobile Users Ting, Kai Ming Multi-dimensional Mass Estimation and Mass-based Clustering Tischer, Peter Accelerating Radius-Margin Parameter Selection for SVMs Using Geometric Bounds Toivonen, Hannu Network Simplification with Minimal Loss of Connectivity







Tong, Hanghang Rare Category Characterization On the Vulnerability of Large Graphs Traina Jr., Caetano QMAS: Querying, Mining and Summarization of Multi-modal Databases Traina, Agma J.M. QMAS: Querying, Mining and Summarization of Multi-modal Databases Tsang, Ivor Location and Scatter Matching for Dataset Shift in Text Mining **Tsourakakis, Charalampos**





On the Vulnerability of Large Graphs



Ueda, Naonori

 Averaged Stochastic Gradient Descent with Feedback: An Accurate, Robust, and Fast Training Method

Vahrenhold, Jan

Resilient K-d Trees: K-Means in Space Revisited

van Leeuwen, Matthijs

■ Subgroup Discovery Meets Bayesian Networks—An Exceptional Model Mining Approach

Venkatachalam, Aravind

☐ A Graph-Based Approach for Multi-folder Email Classification







Vinh, Nguyen Xuan minCEntropy: A Novel Information Theoretic Approach for the Generation of Alternative Clusterings Vlachos, Michail Visualizing Graphs Using Minimum Spanning Dendrograms von Luxburg, Ulrike Multi-agent Random Walks for Local Clustering on Graphs Vuokko, Niko Testing the Significance of Patterns in Data with Cluster Structure Wan, Xiaojun





Multi-document Summarization Using Minimum Distortion



Wang, Chang-Dong A Conscience On-line Learning Approach for Kernel-Based Clustering Wang, Chun-Li Collaborative Learning between Visual Content and Hidden Semantic for Image Retrieval Wang, Dingding Weighted Feature Subset Non-negative Matrix Factorization and Its Applications to Document Understanding Wang, Fan Stratified Sampling for Data Mining on the Deep Web Wang, Fei





Learning a Bi-Stochastic Data Similarity Matrix



□ Compressed Nonnegative Sparse Coding		
Wang, Hao		
□ Causal Discovery from Streaming Features		
Wang, Hua		
☐ A Pairwise-Systematic Microaggregation for Statistical Disclosure Control		
Wang, Jianmin		
☐ Transfer Learning via Cluster Correspondence Inference		
Wang, Jianyong		
 Personalizing Web Page Recommendation via Collaborative Filtering and Topic-Aware Markov Model 		
Wang, Ke		
☐ Enforcing Vocabulary <i>k</i> -Anonymity by Semantic Similarity Based Clustering		







	Anonymizing Temporal Data	
	Probabilistic Inference Protection on Anonymized Data	
	_	
Wang, Peng		
	Efficient Episode Mining with Minimal and Non-overlapping Occurrences	
Wang, Wei		
	K-AP: Generating Specified K Clusters by Efficient Affinity Propagation	
	Efficient Episode Mining with Minimal and Non-overlapping Occurrences	
\٨/	ana Viana	
<u> </u>	ang, Xiang	
	Active Spectral Clustering	
Wang, Xufei		
	Discovering Overlapping Groups in Social Media	







W	ang, Yanqi		
	Frequent Instruction Sequential Pattern Mining in Hardware Sample Data		
Wang, Zheng			
	Homotopy Regularization for Boosting		
Wei, Liang			
	Attribution of Conversion Events to Multi-channel Media		
Weikum, Gerhard			
	Bonsai: Growing Interesting Small Trees		
W	Wells, Jonathan R.		
	Multi-dimensional Mass Estimation and Mass-based Clustering		







Weng, Jianshu

■ What Do People Want in Microblogs? Measuring Interestingness of Hashtags in *Twitter*

Wobcke, Wayne

Learning Collaborative Filtering and Its Application to People to People Recommendation in Social Networks

Wong, Raymond Chi-Wing

- Anonymizing Temporal Data
- Probabilistic Inference Protection on Anonymized Data

Wong, Tak-Lam

Location and Scatter Matching for Dataset Shift in Text Mining







Wong, Weng-Keen Modeling Experts and Novices in Citizen Science Data for Species Distribution Modeling Woznica, Adam Adaptive Distances on Sets of Vectors Wu, Jun Collaborative Learning between Visual Content and Hidden Semantic for Image Retrieval Wu, Junjie Understanding of Internal Clustering Validation Measures Wu, Xindong





10 Years of Data Mining Research: Retrospect and Prospect



SMILE: A Similarity-Based Approach for Multiple Instance Learning NESVM: A Fast Gradient Method for Support Vector Machines Causal Discovery from Streaming Features Wu, Yao Constraint Based Dimension Correlation and Distance Divergence for Clustering **High-Dimensional Data** Xia, Chaolun Efficient Probabilistic Latent Semantic Analysis with Sparsity Control Xiao, Jing Frequent Instruction Sequential Pattern Mining in Hardware Sample Data Xiao, Yanshan Exploiting Local Data Uncertainty to Boost Global Outlier Detection





SMILE: A Similarity-Based Approach for Multiple Instance Learning



Xie, Yan Max-Clique: A Top-Down Graph-Based Approach to Frequent Pattern Mining Xiong, Hui An Unsupervised Approach to Modeling Personalized Contexts of Mobile Users Detecting Blackhole and Volcano Patterns in Directed Networks Understanding of Internal Clustering Validation Measures Xiong, Yuhong D-LDA: A Topic Modeling Approach without Constraint Generation for Semi-defined Classification Xu, Yabo **Anonymizing Temporal Data** Probabilistic Inference Protection on Anonymized Data







Yamada, Keiji

 Exponential Family Tensor Factorization for Missing-Values Prediction and Anomaly Detection

Yan, Jun

A Novel Contrast Co-learning Framework for Generating High Quality Training Data

Yan, Shuicheng

□ A Novel Contrast Co-learning Framework for Generating High Quality Training Data

Yang, Jaewon

Modeling Information Diffusion in Implicit Networks







Yang, Qing An Approach Based on Tree Kernels for Opinion Mining of Online Product Reviews Yang, Qingyan Personalizing Web Page Recommendation via Collaborative Filtering and Topic-Aware Markov Model Yang, Shengwen **Topic Modeling Ensembles** Yang, Zi Term Filtering with Bounded Error Yildirim, Hilmi





SONNET: Efficient Approximate Nearest Neighbor Using Multi-core



Yin, Jie

- Exploiting Local Data Uncertainty to Boost Global Outlier Detection
- SMILE: A Similarity-Based Approach for Multiple Instance Learning

Yu, Hwanjo

□ Passive Sampling for Regression

Yu, Jeffrey Xu

Clustering Large Attributed Graphs: An Efficient Incremental Approach

Yu, Jun

 Modeling Experts and Novices in Citizen Science Data for Species Distribution Modeling







Yu, Kai Training Conditional Random Fields Using Transfer Learning for Gesture Recognition Yu, Kui Causal Discovery from Streaming Features Yu, Lei A Variance Reduction Framework for Stable Feature Selection Yu, Philip S. Multi-label Feature Selection for Graph Classification Exploiting Local Data Uncertainty to Boost Global Outlier Detection Efficient Semi-supervised Spectral Co-clustering with Constraints Transfer Learning on Heterogenous Feature Spaces via Spectral Transformation Probabilistic Inference Protection on Anonymized Data







	Max-Clique: A Top-Down Graph-Based Approach to Frequent Pattern Mining
Yι	ı, Xiao
	Tru-Alarm: Trustworthiness Analysis of Sensor Networks in Cyber-Physical Systems
<u>Y</u> ι	ıan, Yifei
	Scalable Influence Maximization in Social Networks under the Linear Threshold Model
Zł	nang, Changshui
	Homotopy Regularization for Boosting
Zhang, Chongsheng	
	ABS: The Anti Bouncing Model for Usage Data Streams







Zhang, Chunxia An Approach Based on Tree Kernels for Opinion Mining of Online Product Reviews Zhang, Harry An Extensive Empirical Study on Semi-supervised Learning Zhang, Li Scalable Influence Maximization in Social Networks under the Linear Threshold Model Zhang, Ming A Novel Contrast Co-learning Framework for Generating High Quality Training Data **Zhang, Min-Ling** Exploiting Unlabeled Data to Enhance Ensemble Diversity







Zhang, Peng

Classifier and Cluster Ensembles for Mining Concept Drifting Data Streams

Zhang, Xianchao

- Constraint Based Dimension Correlation and Distance Divergence for Clustering High-Dimensional Data
- ☐ Graph-Based Semi-supervised Learning with Adaptive Similarity Estimation

Zhang, Xiangliang

☐ *K*-AP: Generating Specified *K* Clusters by Efficient Affinity Propagation

Zhang, Yanchun

□ A Pairwise-Systematic Microaggregation for Statistical Disclosure Control







Zhang, Yi Training Conditional Random Fields Using Transfer Learning for Gesture Recognition Zhang, Yuan MoodCast: Emotion Prediction via Dynamic Continuous Factor Graph Model **Zhao, Peixiang** gSkeletonClu: Density-Based Network Clustering via Structure-Connected Tree Division or Agglomeration Zheng, Li Hierarchical Ensemble Clustering Zheng, Yaling Active Learning from Multiple Noisy Labelers with Varied Costs







Zheng, Zeyu A Novel Contrast Co-learning Framework for Generating High Quality Training Data **Zhou, Aoying** Detecting Blackhole and Volcano Patterns in Directed Networks Zhou, Fang Network Simplification with Minimal Loss of Connectivity Zhou, Hang Improving Kernel Methods through Complex Data Mapping Zhou, Lizhu Personalizing Web Page Recommendation via Collaborative Filtering and Topic-Aware Markov Model







Zhou, Tianyi NESVM: A Fast Gradient Method for Support Vector Machines Zhou, Yang Clustering Large Attributed Graphs: An Efficient Incremental Approach Zhou, Zhi-Hua Exploiting Unlabeled Data to Enhance Ensemble Diversity Zhu, Huisheng Efficient Episode Mining with Minimal and Non-overlapping Occurrences Zhu, Jun-Yong





A Conscience On-line Learning Approach for Kernel-Based Clustering



Zhu, Qiang Fast and Flexible Multivariate Time Series Subsequence Search Mother Fugger: Mining Historical Manuscripts with Local Color Patches Zhu, Ruixin Transfer Learning on Heterogenous Feature Spaces via Spectral Transformation Zhu, Xingquan Classifier and Cluster Ensembles for Mining Concept Drifting Data Streams Zhuang, Fuzhen D-LDA: A Topic Modeling Approach without Constraint Generation for Semi-defined Classification Zou, Jia





Frequent Instruction Sequential Pattern Mining in Hardware Sample Data



A B C D E F G

HIJKLMN

O P Q R S T U

V W X Y Z





